2018-11-16

The bZIP transcription factor SPA Heterodimerizing Protein represses glutenin synthesis in Triticum aestivum

Plessis, Anne

10.1111/tpj.14163
Plant Journal
Wiley

Article type : Original Article

The bZIP transcription factor SPA Heterodimerizing Protein represses glutenin synthesis in *Triticum aestivum*

Julie Boudet[1*], Marielle Merlino[1*], Anne Plessis[1†], Jean-Charles Gaudin[2‡], Mireille Dardevet[1], Sibille Perrochon[1], David Alvarez[1], Thierry Risacher[3§], Pierre Martre[1¶], Catherine Ravel[1]

[1]UMR GDEC, INRA, Clermont Auvergne University, 63000 Clermont-Ferrand, France
[2]UR BIA, INRA, 44316 Nantes, France
[3]Biogemma, Centre de Recherche de Chappes, 63720 Chappes, France

[†]*Present address: School of Biological Sciences, Plymouth University, Drake Circus, Plymouth PL4 8AA, UK*

[‡]*Present address: UR AGPF, INRA, 45075 Orléans, France*

[§]*Present address: Limagrain Europe, 63720 Chappes, France*

[¶]*Present address: UMR LEPSE, INRA, Montpellier SupAgro, 34060 Montpellier, France*

[*]These authors contributed equally to this manuscript

For correspondence (email: julie.boudet@uca.fr)

Running title: Transcriptional regulation of glutenin synthesis in wheat

Keywords: bZIP transcription factor, gluten, SPA Heterodimerizing Protein, storage proteins, wheat (*Triticum aestivum* L.).

SUMMARY

The quality of wheat grain is mainly determined by the quantity and composition of its grain storage proteins (GSPs). GSPs consist of the low (LMW-GS) and high (HMW-GS) molecular-weight glutenins and gliadins. The synthesis of these proteins is essentially regulated at the transcriptional level and by the availability of nitrogen and sulfur. The regulation network has been extensively studied in barley where BLZ1 and BLZ2, members of the basic leucine zipper (bZIP) family, activate the synthesis of hordeins. To date, in wheat, only the ortholog of BLZ2, Storage Protein Activator (SPA), has been identified as playing a major role in the regulation of GSP synthesis. Here, the ortholog of BLZ1 named SPA Heterodimerizing Protein (SHP) was identified and its involvement in the transcriptional regulation of the genes coding for GSPs was analyzed. In gel mobility shift assays, SHP binds cis-motifs known to bind to bZIP family TFs in HMW-GS and LMW-GS promoters. Moreover, we showed by transient expression assays in wheat endosperms that SHP acts as a repressor of the activity of these gene promoters. This result was confirmed in transgenic lines overexpressing *SHP*, which were grown with low and high nitrogen supply. The phenotype of *SHP* overexpressing lines showed a lower quantity of both LMW-GS and HMW-GS, while the quantity of gliadin was unchanged, whatever the availability in nitrogen. Thus, the gliadin/glutenin ratio was increased, which suggests that gliadin genes may be differently regulated than glutenin genes.

INTRODUCTION

In cereal grains, nitrogen and sulfur to sustain embryo germination and early seedling development are mainly stored in the grain storage proteins (GSPs) gliadin and glutenin. The quantities and proportions of GSPs, which differ in their ability to form polymers, are key determinants of the end-use value of wheat (*Triticum aestivum* L.) grain. Glutenins play an important role in strengthening wheat dough by conferring elasticity, while gliadins

contribute to its viscous properties by conferring extensibility (Branlard *et al.*, 2001). Glutenins can form very large macropolymers during grain desiccation and are composed of low (LMW-GS) and high (HMW-GS) molecular-weight subunits (Shewry *et al.*, 1997; Shewry and Halford, 2002). Gliadins are monomeric proteins that are classified according to their electrophoretic mobility and amino acid sequence as α-, β-, γ- or ω-gliadins.

The quantity and composition of GSPs in mature grain are strongly affected by nitrogen (N) and sulfur (S) nutrition of the parent plant. High N supply increases the amount of GSPs at maturity (Shewry *et al.*, 2001; Triboï *et al.*, 2003; Chope *et al.*, 2014). The GSP subclasses differ in their relative proportions of S-containing amino acids (cysteine and methionine). LMW-GS, α/β- gliadin and γ-gliadin are classified as S-rich, and HMW-GS, ω1,2- gliadin and ω5-gliadin as S-poor (Shewry *et al.*, 1997; Shewry *et al.*, 2001). S-deficiency decreases the concentration of S-rich proteins but it increases the concentration of S-poor proteins with the effect of maintaining a steady level of total GSPs (Zorb *et al.*, 2010; Dai *et al.*, 2015; Bonnot *et al.*, 2017).

During cereal endosperm development, GSP synthesis is mainly controlled at a transcriptional level. The regulatory mechanisms of GSP gene expression in barley have been described as a network of cis-motifs and their interacting TFs (Rubio-Somoza *et al.*, 2006a, b; Moreno-Risueno *et al.*, 2008). This network is conserved in other cereals and dicots as reviewed by Verdier and Thompson (2008) and Xi and Zheng (2011). The bipartite endosperm box has been identified in the promoter of some hordein and LMW-GS genes (Hammond-Kosack *et al.*, 1993; Oñate *et al.*, 1999; Juhász *et al.*, 2011). It contains two distinct protein-binding sites, the GCN4 like motif (GLM, 5'-ATGAG/CTCAT-3') and the prolamin box (P-box, 5'-TGTAAAG-3'), and it plays a key role in activating the expression of GSP genes. The GLM and the P-box are recognized by basic leucine zipper (bZIP) and DNA binding with one finger (DOF) TFs, respectively. The GLM is recognized by BLZ1 and

BLZ2 in barley (Vicente-Carbajosa *et al.*, 1998; Oñate *et al.*, 1999), RISBZ1, REB and RITA-1 in rice (Izawa *et al.*, 1994; Nakase *et al.*, 1997; Onodera *et al.*, 2001), Opaque-2 (O2), OHP1 and OHP2 in maize (Pysh *et al.*, 1993; Zhang *et al.*, 2015) and SPA in wheat (Albani *et al.*, 1997). The P-box is bound by BPBF and SAD in barley and WPBF in wheat (Vicente-Carbajosa *et al.*, 1997; Mena *et al.*, 1998; Diaz *et al.*, 2005). Two additional cis-elements have been described in barley: 5'-AACA/TA-3' binds GAMYB, a TF of the R2R3MYB family, and 5'-TATC/GATA-3' binds HvMCB1 and HvMYBS3, two regulatory proteins of the R1MYB family (Diaz *et al.*, 2002; Rubio-Somoza *et al.*, 2006a, b). Another important motif, the RY box (5'-CATGCATG-3'), is recognized by FUSCA3, a B3-type TF in barley and wheat (Moreno-Risueno *et al.*, 2008; Sun *et al.*, 2017). Differences in the organization of regulatory cis-elements have been noted in wheat GSP promoters. The long endosperm box in the promoter of the LMW-GS gene *GluD3* reported by Hammond-Kosack *et al.* (1993), which contains two copies of the endosperm box, is not present in all LMW-GS promoters (Juhász *et al.*, 2011). The HMW-GS gene promoter contains an atypical endosperm box where the P-box is associated with a G-like box able to bind bZIP proteins like O2 and SPA (Norre *et al.*, 2002; Ravel *et al.*, 2014). Recently a common framework of cis-regulation was found for all HMW-GS gene promoters (Ravel *et al.*, 2014), based on a composite box made of the GATA and GLM motifs (named the GATA-GLM box). This box is functional as it was shown that the GLMs are able to bind SPA. To date, all the TFs characterized as participating in the transcriptional control of GSPs during endosperm development in cereal seeds are activators of the expression of GSP genes as reviewed by Xi and Zheng (2011). In addition to DNA-protein interactions, cooperation between TFs by protein-protein interactions provides an efficient mechanism to control gene expression as reviewed by Xi and Zheng (2011) and Zhang *et al.* (2015).

Members of the bZIP family have been found in fungi, animals and plants (Deppman *et al.*, 2006). In plants, bZIP factors participate in the regulation of light, stress and hormone response pathways, nitrogen/carbon metabolism, pathogen defense, flower development and seed storage and maturation (Jakoby *et al.*, 2002; Schütze *et al.*, 2008; Alonso *et al.*, 2009). They are characterized by a conserved DNA binding domain, which is formed by a region rich in basic amino acids and a leucine zipper that consists of several heptad repeats of hydrophobic residues. This region promotes bZIP homo- or heterodimerization in a specific and predictable manner (Vinson *et al.*, 2002; Deppman *et al.*, 2006). The possible patterns of bZIP factor dimerization are so diverse that many dimer combinations with unique effects on transcription may be generated. In some cases, regulation by bZIPs is made more complex by post-transcriptional regulation like phosphorylation (Schütze *et al.*, 2008) or interactions with other specific TFs or with repressors that modulate subcellular localization of the bZIP proteins. All these regulatory mechanisms are found to be involved in the regulation of cereal GSP gene expression by bZIP proteins. For example, Oñate *et al.* (1999) showed that BLZ2 forms a heterodimer with BLZ1 in yeast. Monomers and dimers of OHP1 or OHP2 are able to bind the O2-like box in the promoter of the 27-kD γ-zein gene (Zhang *et al.*, 2015). Rice REB and RITA-1 proteins are able to form heterodimeric complexes with RISBZ1 (Onodera *et al.*, 2001). The ZmTaxilin protein binds O2 and sequesters it in the cytoplasm, thus negatively modulating O2 activity by preventing it from binding its target genes in the nucleus (Zhang *et al.*, 2012). Ciceri *et al.* (1997) also showed that the binding activity of O2 in maize is regulated diurnally by a reversible phosphorylation mechanism.

The evolution of bZIPs in plants has been extensively studied (Guedes Corrêa *et al.*, 2008; Nijhawan *et al.*, 2008) with a focus on bZIPs homologous to maize O2 (Vincentz *et al.*, 2003). Phylogenetic studies have shown that gene duplication after the separation of monocotyledons and eucotyledons gave rise to two paralogous bZIP groups (Vincentz *et al.*,

2003; Guedes Corrêa *et al.*, 2008). The first group contains O2 orthologs like RISBZ1, BLZ2 and SPA. The second group includes OHP1 orthologs such as rice REB, maize OHP2 and barley BLZ1, which are able to form dimers with the bZIPs of the first group.

Considering O2-related gene evolution (Vincentz *et al.*, 2003) and the ability of O2 orthologs to dimerize with their paralogs strongly suggests that SPA might have a paralog. Sequence homology analysis based on the BLZ1 sequence allowed us to find this TF in the wheat genome. The predicted SHP function in GSP synthesis was studied by *in vitro* and *in vivo* approaches, including overexpressing the *SHP* gene in transgenic lines grown under low and high nitrogen supply. All results point towards SHP acting as a repressor of glutenin synthesis.

RESULTS

SHP is an ortholog of *BLZ1* and is expressed during seed filling

The bZIP transcription factors O2 and OHP in maize and BLZ1 and BLZ2 in barley are able to form homo- and heterodimers and activate transcription by interacting with the GLM in GSP gene promoters. This led us to search for a wheat counterpart that might interact similarly with SPA. The *BLZ1* gene was used in a Blast search on *T. aestivum* genome and three homoeologous sequences on chromosome 5 were retrieved. The CDS of the A, B and D *SHP* copies encoded a polypeptide of 392, 400 and 406 amino acid residues respectively, which include the characteristic DNA binding domain of the bZIP class of TFs, the basic and leucine zipper domains. The translation of these three CDS copies showed their high level of similarity (> 94.5%). Especially, their functional domains are identical. Due to these results, we focussed our work on only one copy. The A copy gene (TraesCS5A02G440400), which was the most expressed in wheat albumen, was further analysed, hereafter designated SPA Heterodimerizing Protein (*SHP*). Multiple sequence

alignment analysis showed that this protein has a high sequence identity with BLZ1 (90.3%), but a lower sequence identity with BLZ2 (32.9%) and SPA (28.8%). The region of highest identity coincides with the characteristic DNA binding domain of the bZIP class of TFs, the basic and leucine zipper domains (Figure 1a). To confirm the evolutionary relationships between SHP and BLZ1, a phylogenetic tree was generated with the deduced amino acid sequences of the bZIP domains of BLZ1, BLZ2, SPA and SHP. SHP is grouped with BLZ1 separate from the paralog group with BLZ2 and SPA (Figure 1b). Relative expression of *SHP* and *SPA* was measured by qRT-PCR in the endosperm of the wheat cultivar NB1 during grain development (Figure 1c). *SHP* relative expression was not changed while *SPA* transcript level increased gradually from 300 to 500 °Cdays after anthesis. *SPA* expression was from two to three fold higher than *SHP* expression during this grain filling phase.

SHP specifically binds the glutenin promoter GLMs and G-box *in vitro*

To test whether the GLMs and G-box of the *GluB1-1* gene promoter are also specifically recognized by SHP, synthetic oligonucleotides containing these motifs were tested by EMSA, where binding of recombinant SHP protein to the DNA sequences is visualized as retarded bands in the gel (Figure 2a). The results revealed that SHP is able to bind the two GLMs and the G-box with different binding affinities. Shifted bands of DNA-protein complexes were clearly observed with the GLM2 and G-box motifs, but that of the GLM1 motif was considerably fainter. The interactions were abolished when the respective motifs were mutated in *glm1*, *glm2* and *G-box* (Figure 2a). The binding specificity of SHP recombinant protein was verified by adding unlabeled intact probes to the reaction, which diminished all retarded bands. SPA can bind the two GLMs and G-box in the promoter of *GluB1-1*, which encodes HMW-GS (Ravel *et al.*, 2014). To evaluate if SHP binds DNA as mono or dimer, we compared the DNA-protein complexes obtained with SPA and SHP and

the GLM2 of the *GluB1-1* gene promoter. Two DNA-protein complexes were observed with SPA and a large shifted band with SHP (Figure S1).

EMSA was also performed with SHP recombinant protein and synthetic oligonucleotides containing the two GLMs from the *GluD3* promoter reported to bind SPA (Albani *et al.*, 1997). The results revealed that SHP is also able to bind these two GLMs with different binding affinities and both interactions were abolished when the motifs were mutated (Figure 2b).

SHP regulates the transcription of *GluB1-1* and *GluD3*

The functional relevance of the *in vitro* interaction between SHP and the GLMs and G-box cis-motifs was investigated *in vivo* by assessing the effect of SHP in transient expression assays in wheat endosperms. Figure 3a shows the constructs used in the assays with the promoter of the HMW-GS *GluB1-1* gene. Immature endosperms were transiently transformed by particle bombardment of these reporters alone or in combination with the pUbi-SHP effector at different molar ratios. Co-transfection of pGluB1-1 with the SHP resulted in respective decreases in GUS activity ($P < 0.0001$) of 23% and 42% for 1:1 and 1:2.5 molar ratios compared with that driven by pGluB1-1 alone (Figure 3b). Based on these results, the same conditions and an equimolar ratio of co-bombarded constructs were used in the following transient expression experiments.

When the reporter construct pGluB1-1*, in which the GLMs are mutated, was transiently expressed, GUS expression was not modified by co-bombardment of the pUbi-SHP effector (Figure 3c). However GUS expression was modified in the presence of SHP when reporter constructs pGluB1-1** (mutated G-box) and pGluB1-1*** (mutated GLMs and G-box) were used. For pGluB1-1** with pUbi-SHP, GUS expression increased ($P < 0.01$) by 36%, while for pGluB1-1*** it decreased ($P < 0.05$) by 15%. Mutation of the two

GLMs, or the G-box, or the GLMs and G-box in the pGluB1-1*, pGluB1-1** and pGluB1-1*** promoters, respectively, resulted in significantly lower basal transcriptional activity with relative decreases of 46%, 58% and 50% respectively compared to the wild-type promoter *GluB1-1* (insert in Figure 3c).

We investigated the functional relevance of the *in vitro* interaction observed between SHP and the GLMs in the LMW-GS *GluD3* promoter by transient expression assays in co-bombarded wheat endosperms. Figure 4a shows the reporter construct used. Developing endosperms were transiently transformed with this reporter alone or in combination with the pUbi-SHP effector at an equimolar ratio. GUS activity induced by pGluD3 alone was very low (Figure 4b). We therefore expressed the transcriptional activator SPA as an effector. GUS activity was not modified by pUbi-SHP (*P = 0.99*) alone, but with pUbi-SPA a threefold increase in GUS activity was induced (Figure 4b). Co-transfection of pGluD3 with pUbi-SPA and pUbi-SHP effectors resulted in a 42% decrease in GUS activity compared with that driven by pGluD3 and pUbi-SPA alone ($P < 0.05$), showing that SHP repressed *GluD3* expression.

SHP expression in _SHP_ overexpressing lines is increased in a nitrogen dependent manner

The regulatory function of *SHP* on GSP synthesis was investigated by producing transgenic wheat lines overexpressing *SHP* (OE). The effect of nutrient availability on GSP quantity and composition was investigated in *SHP* null segregant (NS) and OE lines grown in the greenhouse with low (N-) and high (N+) nitrogen supply. Results from the two *SHP* OE events were similar, therefore they were averaged. Relative expression of *SHP* and *SPA* at 500 °Cdays after anthesis was measured by qRT-PCR (Figure 5). N treatment did not influence the expression of *SHP* in the NS lines. *SHP* expression was five and eight fold

higher in the OE lines compared with the NS lines in the N- and N+ treatments, respectively. In the OE lines, *SHP* expression was 1.5 fold higher in the N+ than in the N- treatment. In the NS lines, *SPA* expression is higher than that of *SHP* (two and three fold in N- and N+ treatments respectively). N treatment influenced the expression of *SPA* in the NS and OE lines. In these latter, *SPA* gene expression was lower than that of *SHP* (more than two fold). *SPA* expression was not influenced by *SHP* overexpression.

Grain storage protein composition in *SHP* overexpressing lines is modified in relation to nitrogen availability

Grain protein concentration and composition in the NS and *SHP* OE lines were determined at maturity under low and high nitrogen supplies. The single grain dry mass, the mass of nitrogen per grain and the grain protein concentration did not differ significantly between the *SHP* OE and NS lines in either nitrogen treatment (Table S1). However, nitrogen supply did have an effect, notably on the grain protein concentration which was 35% and 36% higher in *SHP* OE and NS lines, respectively, under high nitrogen supply.

In both nitrogen treatments, the amount of gliadin per grain was not significantly different between the *SHP* OE and NS lines (Figure 6). However, the amount of glutenin per grain was 9% and 16% lower in the *SHP* OE lines than in the NS lines for the N- and N+ treatments, respectively, although the difference was only statistically significant for the N+ treatment. Consequently, the gliadin-to-glutenin ratios raised significantly in the *SHP* OE lines compared with the NS lines, 20% for the N+ treatment and 15% for N- treatment (Figure 6). The decrease in the quantity of glutenin was due to decreases in the amounts of both LMW-GS and HMW-GS which were significantly lower in *SHP* OE compared with the NS for the N+ treatment (Figure 6). The HMW-GS to LMW-GS ratio was not significantly modified. Quantities of the three classes of gliadin did not differ significantly between the

SHP OE and NS lines in either nitrogen treatments. However, in *SHP* OE we observed weak increases in the quantities of α/β-gliadin (8% in N-) and ω1,2-gliadin (12% in N+), while the quantities of γ-gliadin decreased (- 7% in N+).

DISCUSSION

SPA was identified by Albani *et al*. (1997) who showed that it is an activator of *GluD1-1* in maize and tobacco leaf protoplasts. The orthologs of *SPA* in maize and barley, *O2* and *BLZ2* respectively, activate GSP synthesis (Pysh *et al.*, 1993; Oñate *et al.*, 1999). Their respective paralogs *OHP* and *BLZ1* are also known to enhance the transcription of GSP genes (Vicente-Carbajosa *et al.*, 1998; Zhang *et al.*, 2015). Here we identified and characterized *SHP*. Phylogenetic analysis showed that *SHP* is the ortholog of *BLZ1* and paralogous to the *BLZ2* and *SPA* gene group. Considering the well-described evolution of *O2* related genes (Vincentz *et al.*, 2003), this would be consistent with *SHP* also being the ortholog of *OHP* from maize. TFs involved in GSP regulation described to date are all activators of GSP gene expression. Here we showed that SHP represses glutenin gene expression.

SHP is a repressor of glutenin synthesis independently of nitrogen supply

Like SPA, SHP binds the GLMs and G-box of the *GluB1-1* promoter and the GLMs of the *GluD3* promoter *in vitro*. BLZ1 and OHP proteins have been described as homo-dimerizing when they bind GSP promoter motifs (Vicente-Carbajosa *et al.*, 1998; Zhang *et al.*, 2015). SPA is able to bind the GLM2 of *GluB1-1* promoter. This resulted in two shifted bands of DNA-protein complexes and indicated that SPA is able to bind this cis-motif as monomers and dimers. As recombinant SPA and SHP proteins have similar molecular weight, (47 and 41 kDa, respectively), the large retarded band obtained with SHP also

suggests that this TF could bind as monomers and dimers. Thus SPA and SHP can both bind GLM2 of *GluB1-1* promoter as homodimer and probably other cis-motifs analyzed in this study.

The capacity of barley BLZ1 and maize OHP to activate GSP gene expression prompted us to investigate whether SHP is also involved in the transcriptional regulation of GSPs. Transient experiments in wheat endosperms showed that SHP repressed the *GUS* reporter gene controlled by the *GluB1-1* promoter but not that controlled by the *GluD3* promoter. This difference might be due to the difference in basal expression from each promoter, found to be particularly low for the *GluD3* promoter. To thoroughly investigate SHP binding to the *GluD3* promoter, we adopted a strategy that has already been used to demonstrate the repression activity of TFs in *Arabidopsis* (Thévenin *et al.*, 2012). The latter authors were analyzing *BANYULS* gene (*BAN*) regulation by transient expression assays in *Physcomitrella patens* protoplasts, but decreases in *GFP* reporter gene activity could not be observed. To circumvent this problem, *BAN* promoter activity was increased by a complex composed of three proteins (TT2/AtMYB123, TT8/bHLH042 and TTG1) and decreased by MYBL2, a transcriptional repressor. In wheat endosperms, we used SPA as an activator to increase the basal activity of the *GluD3* promoter, allowing us to ascertain that SHP can significantly repress it. Therefore, SHP appears to be a repressor of glutenin synthesis. This is surprising as its maize and barley orthologs (OHP and BLZ1, respectively) were reported to activate seed storage protein genes. More precisely, OHP activates synthesis of 27-kD γ-zein as confirmed in *OHP* RNAi line which has a dramatic reduction in this GSP (Zhang *et al.*, 2015). Moreover, OHP recognized and trans-activated all of the α-zein promoters, although too much lower levels than did O2 (Yang *et al.*, 2016). In barley, BLZ1 acts as activator of B-hordeins (Vicente-Carbajosa *et al.*, 1998). These opposite results could be explained by the high level of diversity of storage protein genes including promoter regions. Differences exist

in the promoter not only between the genes coding the different classes of seed storage proteins but also between the promoters from genes in a given class suggesting that the regulation of all these genes could differ. As an example, the *in silico* study of the *LMW-GS* gene promoters reported by Juhász *et al.* (2011) showed a high level of polymorphisms in the number and combination of cis-motifs, which could explain the diverse levels of expression of single *LMW-GS* gene. Similarly, distinct α/β gliadin genes show different expression patterns during seed development, which could be explained by differences in the presence of cis-motifs (especially of GLMs) in their promoter sequences (van Herpen *et al.*, 2008; Noma *et al.*, 2016). Therefore, the generalization of the effect of SHP on the expression of all the storage protein genes is likely critical. Some of these genes might be activated while others repressed. However, the phenotype of OE lines confirmed the repressive activity of SHP on glutenins. The LMW-GS and HMW-GS quantities in grain of transgenic lines were significantly decreased to a similar extent, such that the HMW-GS to LMW-GS ratio was not different in the SHP OE lines compared to the NS lines. No significant effect of the overexpression of *SHP* on the gliadin fractions was observed. Thus our work confirms the role of SHP in the regulation of storage proteins.

The effect of *SHP* overexpression on GSP composition is independent of nitrogen availability as similar changes in GSP composition occurred under the two nitrogen treatments. Assuming mRNA and protein abundance correlate as shown for wheat GSPs (Dai *et al.*, 2015), the higher expression of *SHP* in the OE lines when nitrogen was available did not seem to have a larger impact. In the OE lines, *SHP* was under the control of a HMW-GS promoter which responds strongly to nitrogen (Dai *et al.*, 2015). The GLM motif is known to play an important role in the transcriptional response of GSP genes to nitrogen. Indeed, it was found to be essential for the activation of GSP gene transcription in response to amino acids and ammonium (Muller and Knudsen, 1993). While the N-response is mediated by GLM, an

O2 binding site, it does not require O2 protein to take effect (Muller *et al.*, 1997). In the *O2* homozygous mutant, 22-kDa zeins are greatly reduced but are synthetized in response to nitrogen supply only when O2-binding sites are intact. Our results strongly suggest that similarly to O2 in maize, SHP does not mediate the N-response of GSP genes in wheat.

SHP is involved in a complex regulatory mechanism

The regulatory mechanisms of GSP gene expression in barley have been described as a network of DNA-protein and protein-protein interactions (Rubio-Somoza *et al.*, 2006a, b; Moreno-Risueno *et al.*, 2008). As SPA is known to bind the GLMs and G-box cis-motifs of the *GluB1-1* promoter (Ravel *et al.*, 2014), transcriptional activity of SPA on the *GluB1-1* gene was analyzed (Figure S2). Co-transfection of pGluB1-1 and the SPA effector resulted in a significant increase in GUS activity compared with that driven by the pGluB1-1 alone ($P < 0.05$). SPA and SHP, which compete for the same binding site, act respectively as activator and repressor TFs. Gene-specific repression was often thought to either direct or indirect transcriptional repression mechanism (Gaston and Jayaraman, 2002). Indirect or passive repression mechanisms can be thought by competition between an activator and a repressor for a common binding site. If mRNA and protein abundance are correlated, then SPA is more abundant than SHP in grain (two or three fold higher than SHP). Thus SPA may bind GLMs and the G-box to activate transcription for a major *GluB1-1* promoter activity (Figure 7). On the contrary, SHP is more abundant than SPA in OE lines or in endosperm transiently transformed with the pUbi-SHP. Thus, SHP binds all bZIP cis-motifs, resulting in a significant decrease in *GluB1-1* activity. SHP may prevent SPA binding and its activating activity (Figure 7).

Nevertheless, the passive repression of SHP could not explain all observations. The analysis of the effect of SHP on mutated *GluB1-1* promoter activity could suggest other mechanisms for SHP action. When the G-box is mutated, a minor activation of the *GluB1-1* promoter is observed. Thus, in this condition, SHP acts as a minor activator. The function of bZIP TFs depends strictly on their ability to dimerize (Llorca *et al.*, 2015). While interacting bZIP monomers have varying transactivation and DNA-binding properties, homo- or heterodimerization generates diverse functions (Weltmeier *et al.*, 2006; Llorca *et al.*, 2015). These bZIP properties and our results could suggest that SHP-SPA heterodimers may be involved in *GluB1-1* regulation.

Transcriptional regulation is exerted by the concerted action of multiple TFs and a specific conformation of DNA that allows strong activation of gene expression. Additive effect of other TFs could be considered for *GluB1-1* regulation. Recently, Sun *et al.* (2017) identified TaFUSCA3, a B3 TF, which binds an RY motif in the *GluB1-1* promoter, activating the promoter and interacting with SPA protein. The TaGamyb binds AACA motif and activates *GluD1-2* gene promoter in *Arabidopsis* (Guo *et al.*, 2015). Near the GLM-GATA box, AACA and RY motifs were identified (Ravel *et al.*, 2014). In barley, HvMYBS3, which binds GATA motif, enhances the GUS expression levels controlled by the *Itr1* promoter (Rubio-Somoza *et al.*, 2006a). Moreover, HvMYBS3 can form a yeast ternary complex with the binary complex of BPBF and BLZ2. Major activity of the *GluB1-1* gene promoter is observed when the GLMs of the GATA-GLM box and G-box are intact. This observation is consistent with previous results (Norre *et al.*, 2002; Ravel *et al.*, 2014). Therefore protein-protein interaction, protein complexes formation are probably involved in *GluB1-1* regulation. According to the annotation of HMW-GS promoters, the cis-motif organization is conserved and suggesting that all HMW-GS genes are regulated by the same mechanisms (Ravel *et al.*, 2014).

In conclusion, we identified and characterized *SHP*, the wheat ortholog of *BLZ1* and *OHP*. We demonstrate that SHP regulates glutenin synthesis. SHP represses glutenin gene expression while BLZ1 and OHP activate B-hordein and zein expression, respectively. The results suggest different possible mechanisms for SHP activity on GSP regulation: passive repressive activity, bZIP dimerization and protein-protein interactions. Further analyses are needed to prove whether SPA and SHP TFs interact and to explore the possible role of additional TFs in this regulatory network. The quantity and composition of GSPs are the main determinants of the rheological and bread-making properties of wheat dough. Therefore it will be interesting to study the natural variability of *SHP* expression to modulate the gliadin/glutenin ratio.

EXPERIMENTAL PROCEDURES

Identification and bioinformatics analysis of the *SHP* gene

To identify the putative *SHP* gene in wheat (*Triticum aestivum*), the sequence of the *BLZ1* gene (GenBank: X80068.1) from *Hordeum vulgare* was used as a query probe in a Blast search of the wheat data library (http://plants.ensemble.org/Triticum_aestivum/Info/Index). Three homoeologous *T. aestivum* gene sequences with high identity to the ORF of *BLZ1* were found. A phylogenetic analysis was performed to compare the deduced amino acid sequence of the conserved bZIP domain of SHP with those of BLZ1 and BLZ2 (GenBank: CAA71795.1) from barley and SPA (GenBank: CAA70216.1) from *T. aestivum*. The alignment of the bZIP domains was performed by using the software programs MEGA 7 (Kumar *et al.*, 2016) and ClustalW (Thompson *et al.*, 1994) with default parameters. The UPGMA method based on a JTT matrix-based model was used to construct a phylogenetic tree. The rate of variation among

sites was modeled with a gamma distribution (shape parameter = 4) and 1,000 bootstrap samplings were made.

Total RNA extracted from developing wheat seed (300 °Cdays after anthesis) of the wheat cultivar Courtot was used to synthesize first-strand cDNA to then amplify the cDNA of *SHP*. This gene was amplified by the use of forward (5'- GTCCCCCGGCGTATTCTC-3') and reverse (5'- CTGCCCAACAATAATTTCA-3') primers and a touchdown PCR program (annealing temperatures decreasing from 65 °C to 55 °C). The PCR product was purified and cloned into pGEM-T Easy vector (Promega), which was sequenced to confirm the *SHP* cDNA sequence had been amplified.

Cloning, expression and purification of recombinant SHP protein in *Escherichia coli*

To produce a recombinant SHP protein, the full-length cDNA was amplified by PCR using primers containing restriction sites: *Bam*HI for the forward primer (5'-NNNNGGATCCGAGCGCGTCTTCTCCGTCG-3') and *Hind*III for the reverse primer (5'-CTATGAGGTCGATCCGGAAAGCTTNNNN-3'). *SHP* cDNA was inserted into the pET32-TEV plasmid (Novagen, Merck) at the *Bam*HI and *Hind*III sites just downstream of the sequence encoding the TEV protease cleavage site ENLYFQ/G, cleavage occurring between Q and G. The recombinant protein thus produced would consist of an N-terminal thioredoxin (Trx) fused with SHP (Trx-SHP) and six histidine residues (HisTag) upstream from the TEV protease cleavage site. Cleavage by the TEV protease makes it possible to recover recombinant SHP identical to wild-type SHP except that the starting M is substituted by the dipeptide GS. The SHP protein was expressed in *E. coli* BL21-DE3 strain (Invitrogen, Life Technologies). Bacteria were grown in LB medium (50 µg mL^{-1} ampicillin) at 37 °C. When cells reached an optical density of 0.6 the expression of the recombinant protein was induced by addition of 0.5 mM IPTG (isopropyl β-D-1-thiogalactopyranoside) overnight at

28 °C. After harvesting, pelleted cells were resuspended in lysis buffer (20 mM Tris-HCl pH 8.0, 200 mM NaCl, 20 mM imidazole, 0.5 mg mL^{-1} lysozyme, 5 µg mL^{-1} DNAse I and 10 µg mL^{-1} RNAse) then sonicated 6 times for 30 s each.

The soluble fraction was purified using a HisTrap TMFF column (GE Healthcare, Vélizy, France) for Ni^{2+} affinity chromatography. Elution was performed with a 50 to 300 mM imidazole gradient. This first purification step was followed by anion exchange chromatography with Resource Q resin (GE Healthcare) on an Äkta Avant system (GE Healthcare). The purity (> 95%) was assessed by SDS-PAGE and the protein concentration was determined with the Pierce bicinchoninic acid (BCA) protein assay kit (Fisher Scientific, Illkirch, France) using BSA as a standard. Protein identity was checked by western blot and mass spectrometry. Trx-SHP was directly cleaved in a reaction supplemented with 1 mM DTT with TEV protease (Sigma) at an enzyme:substrate ratio of 1:100, incubated at 30 °C for 2 h. Digested protein was applied to a HisTrap TMFF column in order to remove Trx via the HisTag. SHP in the flow-through was applied to an anion exchange chromatography column with Resource Q resin and eluted to separate it from TEV protease. The purity and integrity of SHP recombinant protein was assessed by SDS-PAGE. Before EMSA experiments, the protein was dialyzed against a buffer containing 10 mM HEPES pH 7.4, 150 mM NaCl and 0.05% (v/v) Tween-20. The dialysate was then concentrated with an Amicon 10 kDa filter (Millipore).

Electrophoretic mobility shift assay (EMSA)

The probes containing the putative consensus bZIP binding sites from the *GluB1-1* (GLM1, GLM2 and G-box annotated by Ravel *et al*., 2014) and *GluD3* (GLM1 and GLM2 described in Albani *et al.*, 1997) gene promoters and their mutated versions were synthesized by Sigma (sequences provided in Figure 2). The ssDNA probes were labelled and hybridized

as described in Ravel *et al.* (2014) using the biotin 3' End DNA Labeling Kit (Pierce). The labelled dsDNA probes (20 fmol) were incubated with the cleaved SHP protein (800 ng) for 30 min at room temperature in 20 µL of binding buffer containing 10 mM Tris pH 7.5, 2 mM dithiothreitol, 100 mM KCl, 10% (v/v) glycerol, 0.05% (v/v) nonyl phenoxypolyethoxylethanol, 2 mM ethylenediaminetetraacetic acid, 100 ng μL^{-1} poly(dI.dC), and 0.2 µL of protease inhibitor cocktail (P9599, Sigma). When required, unlabelled double-stranded oligonucleotides were included immediately prior to probe addition in the excess amounts (400× the amount of unlabeled GLM1 and G-box probes or 200× the amount of unlabeled GLM2 probe). DNA-protein complexes were analyzed by non-denaturing 5% (w/v) polyacrylamide gel electrophoresis as in Ravel *et al.* (2014).

DNA construct for transient expression assays

The promoter of the *GluB1-1* gene (termed GluB1-1), encoding the Bx7 subunit, previously analyzed by Ravel et al. (2014) was used for particle bombardment. In addition, to assess the role of the GLMs and G-box motifs, the same fragment of promoter with mutations in the two GLMs (termed GluB1-1*), or in the G-box (termed GluB1-1**), or in both GLMs and the G-box (termed GluB1-1***) were synthesized (Figure 3a). The *GluD3* gene promoter, a 346 bp fragment upstream of the start codon containing two GLMs (termed GluD3) was also synthesized (Figure 4a). The complete *SHP* and *SPA* cDNAs from the A genome, each of them under the control of the maize *Ubiquitin* promoter plus the first intron of the *Ubiquitin* gene were used as effector constructs. All constructs used for transient expression assays were obtained using Gateway technology (Invitrogen). Three entry clones were used (pDONRP4-P1R, pDONR221 and pDONRP2R-P3) to obtain the expression vector pDESTR4-R3. pDONRP4-P1R contained the maize *Ubiquitin* promoter plus the first intron of the *Ubiquitin* gene. pDONR221 contained the reporter genes (*GUS* or *GFP*) or the

ORF (*SHP* or *SPA*). pDONRP2R-P3 contained the 3'-terminator *nopaline synthase* gene (3'-NOS). The pDESTR4-R3 based expression vectors (pGluB1-1, pGluB1-1*, pGluB1-1**, pGluB1-1***, pGluD3 with *GUS* gene reporter, pUbi-GFP, pUbi-SHP and pUbi-SPA) were thus created through Gateway combination.

Endosperms from cv. Recital were collected at 230 °Cdays after anthesis from plants grown in a controlled culture chamber averaging 19 °C per day. Plants received 68 mL column^{-1} day^{-1} of water or nutrient solution. For four weeks, plants received a 4 mmol N L^{-1} nutrient solution (N4) containing 1 mM KH_2PO_4, 1 mM KNO_3, 0.5 mM $Ca(NO_3)_2$, 0.5 mM NH_4NO_3, 0.1 mM $MgSO_4$, 2 mM $MgCl_2$, 3.5 mM $CaCl_2$, 4 mM KCl for macroelements, and 10 µM H_3BO_3, 0.7 µM $ZnCl_2$, 0.4 µM $CuCl_2$, 4.5 µM $MnCl_2$, 0.22 µM MoO_3, and 50 µM EDFS-Fe for microelements. Then, after earing, plants received water. Gold particle coating and bombardment were performed according to Ravel *et al.* (2014). After bombardment, endosperms were incubated for 24 h in the dark at 30 °C in a Murashige and Skoog medium supplemented with 3% (w/v) sucrose. *GUS* and *GFP* expression was quantified according to Ravel *et al.* (2014). The pUbi-GFP construct was used to determine the efficiency of bombardment and to calculate the normalized *GUS* expression (the number of GUS foci divided by the number of GFP foci). For each combination of constructs, three to eight independent bombardments of three Petri dishes containing eight endosperms each were performed.

Production of *SHP* transgenic lines and growth conditions

Immature seeds of the spring wheat (*Triticum aestivum*) line NB1 were transformed by *in planta* inoculation using *Agrobacterium tumefaciens* and transgenic lines were regenerated as explained in Risacher *et al.* (2009). The full-length *SHP* cDNA sequence was previously amplified by PCR from a cDNA library extracted from immature seeds of the

bread wheat cultivar Courtot. The vector pSB11 was used to produce transgenic plants expressing sense *SHP-A* cDNA controlled by the promoter of *GluD1-1*, a HMW-GS gene encoding the Dx5 subunit (Lamacchia *et al.*, 2001), and the Nos terminator (Figure S3). The plasmid includes the kanamycin resistance cassette nptII controlled by the *Actin* promoter and Nos terminator for selection. The resulting pSB11-based plasmid was introduced into the LBA4404 (pSB1) Agrobacterium strain where it recombined to form a superbinary vector (Komari *et al.*, 1996). For each transformation event, the number of T-DNA insertions was evaluated by Southern blot and transformants with several copies of the transgene were discarded. The integrity of the transgene in the lines with a single insertion was verified by PCR. Lines representing ten transformation events with a unique copy of the transgene were obtained. For each of these, T1 plants were self-pollinated to generate the T2 generation composed of 25% homozygotes, 50% hemizygotes and 25% null segregant plants (NS). The zygosity of progenies from self-pollinated homozygotes and their respective NS was checked by qPCR. Selfing of confirmed homozygote and NS T2 plants gave rise to the T3 generation, i.e. the OE line and NS to be used as controls. *SHP* overexpression was measured by qPCR from RNA extracted from seeds at 400 °Cdays after anthesis. Lines derived from the two transformation events giving rise to the highest level of overexpression were further studied.

To study the effects of *SHP* overexpression, plants representing each OE transformation event and a mix in equal proportion of their respective NS were grown in a greenhouse with two levels of nitrogen supply. T4 seeds were germinated for two to three days at room temperature on wet filter paper in Petri dishes. Germinated seeds were then transferred to soil in 50-mL PVC columns (inner diameter 7.5 cm, length 50 cm, 2 plants per column) and arranged in the greenhouse in a strip-plot design with the genotypes as rows and the N treatments as columns with four replicated blocks to form a homogeneous stand with a plant density of 261 plants m^{-2} (that is, at a similar plant density as in the field). Temperature

was controlled at 22 °C during the day and 18 °C during the night. Day length was 16 h supplemented with artificial light when needed. Plants received 68 mL column^{-1} day^{-1} of water or nutrient solution. For four weeks plants received a 6 mmol N L^{-1} nutrient solution (N6) containing 1 mM KH_2PO_4, 2 mM KNO_3, 1.5 mM $Ca(NO_3)_2$, 0.5 mM NH_4NO_3, 0.2 mM $MgSO_4$, 1.8 mM $MgCl_2$, 1.5 mM $CaCl_2$, 2 mM KCl for macroelements, and 10 µM H_3BO_3, 0.7 µM $ZnCl_2$, 0.4 µM $CuCl_2$, 4.5 µM $MnCl_2$, 0.22 µM MoO_3, and 50 µM EDFS-Fe for microelements. Then until anthesis plants received a 3 mmol N L^{-1} nutrient solution (N3), which was the same as the N6 solution except the macroelement component was diluted two-fold. At anthesis, continuous water irrigation was used to remove any excess nutrient solution from the soil; then irrigation with water was maintained at its previous level. At 300 °Cdays after anthesis, the columns were rinsed again, then received either the N12 nutrient solution, which was the same as N6 except it contained 4 mM KNO_3, 3 mM $Ca(NO_3)_2$, 1 mM NH_4NO_3, 2 mM $MgCl_2$, no $MgSO_4$, no $CaCl_2$ and no KCl, or a nutrient solution containing no nitrogen (N6 solution with no KNO_3, no $Ca(NO_3)_2$, no NH_4NO_3 and no $MgSO_4$ but with 2 mM $MgCl_2$, 3 mM $CaCl_2$ and 4 mM KCl) until grain ripeness. Main-stem ears were tagged when the anthers of the central florets appeared. Degree-days (°Cdays) were calculated as the sum of the average daily temperatures after anthesis with a base temperature of 0 °C.

Determination of total protein concentration and storage protein composition of SHP transgenic lines

Grains were sampled at maturity 1050 °Cdays after anthesis. For each treatment, four main-stem ears were sampled. Four biological replicates were used corresponding to the four replicated blocks. Grain dry mass and total nitrogen concentration of a sub-sample of grains (ca. 65%) were measured. The remaining grains were lyophilized to calculate the percentage of remaining water.

Grains were milled for 2 min using a custom ball mill. An aliquot of 5 mg of wholemeal flour was weighed in tin capsules and the total N concentration was determined with the Dumas combustion method (Association of Analytical Communities International approved method no. 992.23) using a FlashEA 1112 N/Protein Analyzer (Thermo Electron Corp, Waltham, MA). Grain protein concentration (GPC) was calculated by multiplying grain N concentration by 5.62 (Mosse *et al.*, 1985).

Non-prolamin and gliadin protein fractions were sequentially extracted from 66.6 mg of wholemeal flour as described by Triboi *et al.* (2003) and modified by Plessis *et al.* (2013). Each 2-ml tube contained one stainless steel bead (5 mm diameter) and samples were stirred by placing the tubes on a rotating wheel (40 rpm) during each extraction and washing step. The non-prolamin protein fraction was extracted for 30 min at 4 °C with 1 ml of 50 mM phosphate buffer (pH 7.8) containing 0.1 M NaCl. After centrifugation for 10 min (18 000 g) at 4 °C, the supernatant was collected and the pellet was washed twice for 10 min each with 1 ml of the same buffer. After centrifugation in the same conditions, all supernatants were pooled. The same steps were used to extract the gliadin protein fraction from the previous pellet with 70% (v/v) ethanol. The supernatant (80 µl) of non-prolamin protein fraction was oven-dried overnight at 60 °C in tin capsules and its total N concentration was determined with the Dumas combustion method as described above.

The gliadin extracts used were those obtained by sequential extraction, but glutenins were extracted independently from 100 mg of flour with a 25mM borate buffer pH8, 50% (v/v) propanol-1, 1% (w/v) DTT adapted from Fu and Kovacs (1999). Gliadin classes (ω1,2-, α/β-, and γ-gliadin) and glutenin sub-units (HMW-GS and LMW-GS) were separated and quantified by HPLC using an Agilent 1290 Infinity LC system (Agilent Technologies, Santa Clara, CA, http://www.agilent.com) as described in Dai *et al.* (2015). The gliadin and glutenin extracts were filtered through regenerated cellulose syringe filters (0.45-µm pore

diameter, UptiDisc; Interchim, http://www.interchim.com), and 4 µL of gliadin or 2 µL of glutenin extract was injected into a C8 reversed-phase Zorbax 300 StableBond column (2.1 × 100 mm, 3.5 µm, 300 A; Agilent Technologies) maintained at 50 °C. The eluents used were ultra-pure water (solvent A) and acetonitrile (solvent B), each containing 0.1% trifluoroacetic acid. The flow rate was 1 mL min^{-1}. Proteins were separated by using a linear gradient, from 24% to 50% solvent B over 13 min for gliadin, and from 23% to 42% solvent B over 25 min for glutenin. Proteins were detected by UV absorbance at 214 nm. Chromatograms were processed with CHEMSTATION 10.1 software (Agilent Technologies). The signal obtained from a blank injection was subtracted from the chromatograms before integrating the data. The HPLC peaks corresponding to each of the three gliadin classes were identified following the observations of Wieser *et al.* (1998). A calibration curve established by the quantification of a same standard by RP-HPLC and Dumas analysis was used to calculate the quantity of each gliadin class or glutenin subunit. By multiplying the quantity of each protein measured in the dry flour by the grain dry mass, the quantity of each gliadin class or glutenin subunit per grain was obtained.

SHP and _SPA_ expression measurements in cv. NB1 and SHP transgenic lines

Four grains per ear were sampled for RNA analysis, the embryos were cut out and the rest of the grain immediately frozen in liquid nitrogen and kept at -80 °C. *SHP* and *SPA* expressions were quantified in transgenic lines at 500 °Cdays after anthesis during the linear phase of starch and protein accumulation. Kinetics of *SHP* and *SPA* expression in cv. NB1 was made in a previous experiment during grain development from 300 to 700 °Cdays. RNA of cv. NB1 and transgenic grains was extracted from 75 mg of grain powder in 750 µL of extraction buffer (200 mM Tris-HCl pH 9, 400 mM KCl, 200 mM sucrose, 35 mM $MgCl_2$, 25 mM EDTA) and 600 µL phenol/chloroform pH 8. The suspension was homogenized by

vortexing for 30 s and then centrifuged for 10 min at 15,000 × *g*. The supernatant was collected. The pellet was resuspended in 600 µL of phenol/chloroform, centrifuged in the same conditions and the supernatant collected, and the whole step repeated. Supernatants were pooled. RNA was precipitated by adding 1 M acetic acid (0.1 volume) and ethanol (2.5 volumes). The RNA pellet was washed with 3 M Na acetate (pH 6) and resuspended in water. A second acetic acid/ethanol precipitation was performed before resuspending the pellet in 50 µL RNase free water. RNA was treated with RNase-free DNase according to the supplier's instructions (AMBION). RNA in solution was quantified by measuring the absorbance at 260 nm in a spectrophotometer. Approximately 2 µg of total RNA was reverse transcribed using oligo(dT)20 and reverse transcriptase (Bio-Rad iScriptTM Select cDNA Synthesis kit) in a final volume of 40 µl. Transcript levels of three housekeeping genes and TFs genes were quantified by real-time q-PCR using Lightcycler 480 SYBR Green I Master (Roche) in 15 µl reactions with 5 µl of cDNA diluted 10 fold. Relative expression (RE) was calculated as: $RE = \varepsilon^{Cp} / 2^{Cp^*}$, where ε is the efficiency of the primers for the measured gene, Cp is the Cp for the measured gene and Cp* the geometric mean of the housekeeping genes Cp (Pfaffl *et al.*, 2004). The efficiency of the primers for the housekeeping genes was closed to 2. The normalization was made with regard to the most stable three housekeeping genes in the chosen experimental conditions using the geNorm algorithm (Vandesompele *et al.*, 2002). The primer sequences are given in Supplementary Table S2.

Statistical analysis

Results of the transient expression assays were analyzed by ANOVA with promoter construct as the factor and normalized GUS expression as the variable followed by a post-hoc Dunnett test. Each mean value was compared to that of the reporter construct without

effector. Tukey's test was used to compare the means of two normalized GUS expression values obtained with different combinations of reporters and effectors.

Differences in total grain protein concentration and percentages of LMW-GS, HMW-GS and gliadins in total grain N were analyzed by using an ANOVA with two factors, N treatment and genotype, followed by a post-hoc Dunnett test to compare the mean due to each transgene insertion event to that of the NS in each treatment. For each trait, as the two *SHP* OE events behaved similarly, we grouped data from both lines and thus considered two points per block (i.e. means and standard error were calculated from 8 datasets). Statistical differences were judged at the 0.05 confidence level.

ACKNOWLEDGEMENTS

We thank the staff of the UMR GDEC experimental platform (INRA, Clermont Auvergne University) for help in the greenhouse and Bénédicte Bakan (UR BIA, INRA, Nantes) for production of the pET32-TEV plasmid. We thank Laurence Lavenant and Alexandre Gay for their technical assistance from the Recombinant Proteins Lab (UR BIA, INRA, Nantes). This work was supported by the French Government managed by the National Research Agency (ANR) and France AgriMer in the framework of Investments for the Future BreedWheat project (ANR-10-BTBR-03).

The authors declare no conflicts of interest.

SUPPORTING INFORMATION

Table S1. Single grain dry mass, total quantity of N per grain and grain protein concentration at maturity (1050 °Cdays) from *SHP* overexpressing (OE) and the corresponding null segregants (NS) wheat lines.

Table S2. Sequences of the primers used for qRT-PCR.

Figure S1. Electrophoretic mobility shift assays (EMSA) of the recombinant SHP and SPA proteins with GLM2 derived from the *GluB1-1* gene promoter.

Figure S2. Transient expression assays of *GluB1-1* promoter activity with SPA in developing wheat endosperms.

Figure S3. *SHP* transgene construct used for wheat transformation.

REFERENCES

Albani D., Hammond-Kosack M.C.U., Smith C., Conlan S., Colot V., Holdsworth M. and Bevan M.W. (1997) The wheat Transcriptional Activator SPA: a seed-specific bZIP protein that recognizes the GCN4-like motif in the bifactorial endosperm box of prolamin genes. *Plant Cell* **9**, 171-184.

Alonso R., Oñate-Sánchez L., Weltmeier F., Ehlert A., Diaz I., Dietrich K., Vicente-Carbajosa J. and Dröge-Laser W. (2009) Pivotal role of the Basic Leucine Zipper transcription factor bZIP53 in the regulation of *Arabidopsis* Seed maturation gene expression based on Heterodimerization and protein complex formation. *Plant Cell* **21**, 1747–1761.

Bonnot T., Bancel E., Alvarez D., Davanture M., Boudet J., Pailloux M., Zivy M., Ravel C. and Martre P. (2017) Grain subproteome responses to nitrogen and sulfur supply in diploid wheat *Triticum monococcum* ssp. *monococcum*. *Plant J.* **91**: 894-910

Branlard G., Dardevet M., Saccomano R., Lagoutte F. and Gourdon J. (2001) Genetic diversity of wheat storage proteins and bread wheat quality. *Euphytica* **119**, 59–67.

Chope G.A., Wan Y., Penson S.P., Bhandari D.G., Powers S.J., Shewry P.R. and Hawkesford M.J. (2014) Effects of genotype, season, and nitrogen nutrition on gene expression and protein accumulation in wheat grain. *J. Agric. Food Chem.* **62**, 4399–4407.

Ciceri P., Gianazza E., Lazzari B., Lippoli G., Genga A., Hoscheck G., Schmidt R.J. and Viotti A. (1997) Phosphorylation of Opaque2 changes diurnally and impacts its DNA binding activity. *Plant Cell* **9**, 97-108.

Dai Z., Plessis A., Vincent J., Duchateau N., Besson A., Dardevet M., Prodhomme D., Gibon Y., Hilbert G., Pailloux M., Ravel C. and Martre P. (2015) Transcriptional and metabolic alternations rebalance wheat grain storage protein accumulation under variable nitrogen and sulfur supply. *Plant J.* **83**, 326–343.

Deppman C.D., Alvania R.S. and Taparowsky E.J. (2006) Cross-species annotation of basic leucine zipper factor interactions: insight into the evolution of closed interaction networks. *Mol. Biol. Evol.* **23**, 1480-1492.

Diaz I., Vicente-Carbajosa J., Abraham Z., Martinez M., Isabel-LaMoneda I. and Carbonero P. (2002) The GAMYB protein from barley interacts with the DOF transcription factor BPBF and activates endosperm-specific genes during seed development. *Plant J.* **29**, 453-464.

Diaz I., Martinez M., Isabel-LaMoneda I., Rubio-Somoza I. and Carbonero P. (2005) The DOF protein, SAD, interacts with GAMYB in plant nuclei and activates transcription of endosperm-specific genes during barley seed development. *Plant J.* **42**, 652-662.

Fu B.X. and Kovacs M.I.P. (1999) Rapid single-step procedure for isolating total glutenin proteins of wheat flour. *J. Cereal Sci.* **29**, 113–116.

Gaston G. and Jayaraman P.-S. (2002) Transcriptional repression in eukaryotes: repressors and repression mechanisms. *Cell. Mol. Life Sci.* **60**, 721-741.

Guedes Corrêa L.G., Riaño-Pachón D.M., Guerra Schrago C., Vicentini dos Santos R., Mueller-Roeber B. and Vincentz M. (2008) The Role of bZIP transcription factors in green plant evolution: adaptive features emerging from four founder genes. *PLoS ONE* **3**, e2944.

Guo W., Yang H., Liu Y., Gao Y., Ni Z., Peng H., Xin M., Hu Z., Sun O. and Yao Y. (2015) The wheat transcription factor TaGAMyb recruits histone acetyltransferase and activates the expression of a high-molecular-weight glutenin subunit gene. *Plant J.* **84**, 347-359.

Hammond-Kosack M.C.U., Holdsworth M. and Bevan M.W. (1993) *In vivo* footprinting of a low molecular weight glutenin gene (*LMWG-1D1*) in wheat endosperm. *EMBO J.* **12**, 545-554.

Izawa T., Foster R., Nakajima M., Shimamoto K. and Chua N.H. (1994) The rice bZlP ranscriptional activator RlTA-1 is highly expressed during seed development. *Plant Cell* **6**, 1277-1287.

Jakoby M., Weisshaar B., Dröge-Laser W., Vicente-Carbajosa J., Tiedemann J., Kroj T. and Parcy F. (2002) bZIP transcription factor in *Arabidopsis*. *Trends Plant Sci.* **7**, 106-111.

Juhász A., Makai S., Sebestyén E., Tamás L. and Balázs E. (2011) Role of conserved non-coding regulatory elements in LMW glutenin gene expression. *PLoS ONE* **6**, e29501.

Komari T., Hiei Y., Saito Y., Murai N. and Kumashiro T. (1996) Vectors carrying two separate T-DNAs for co-transformation of higher plants mediated by Agrobacterium tumefaciens and segregation of transformants free from selection markers. *Plant J.* **10**, 165–174.

Kumar S., Stecher G. and Tamura K. (2016) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870-1874.

Lamacchia C., Shewry P.R., Di Fonzo N., Forsyth J.L., Harris N., Lazzeri P.A., Napier J.A., Halford N.G. and Barcelo P. (2001) Endosperm-specific activity of a storage protein gene promoter in transgenic wheat seed. *J. Exp. Bot.* **52**, 243–250.

Llorca C.M., Berendzen K.W., Malik W.A., Mahn S., Piepho H.P. and Zentgraf U. (2015) The elucidation of the interactome of 16 Arabidopsis bZIP factors reveals three independent functional networks. *PLoS ONE* **10**, e0139884.

Mena M., Vicente-Carbajosa J., Schmidt R.J. and Carbonero P. (1998) An endosperm-specific DOF protein from barley, highly conserved in wheat, binds to and activates transcription from the prolamin-box of a native b-hordein promoter in barley endosperm. *Plant J.* **16**, 53–62.

Moreno-Risueno M.A., Gonzalez N., Diaz I., Parcy F., Carbonero P. and Vicente-Carbajosa J. (2008) FUSCA3 from barley unveils a common transcriptional regulation of seed-specific genes between cereals and *Arabidopsis*. *Plant J.* **53**, 882–894.

Mosse J., Huet J. and Baudet J. (1985) The amino-acid composition of wheat-grain as a function of nitrogen-content. *J. Cereal Sci.* **3**, 115–130.

Muller M. and Knudsen S. (1993) The nitrogen response of a barley C-Hordein promoter is controlled by positive and negative regulation of the Gcn4 and endosperm Box. *Plant J.* **4**, 343–355.

Muller M., Dues G., Balconi C., Salamini F. and Thompson R.D. (1997) Nitrogen and hormonal responsiveness of the 22 kDa α-zein and b-32 genes in maize endosperm is displayed in the absence of the transcriptional regulator Opaque-2. *Plant J.* **12**, 281-291.

Nakase M., Aoki N., Matsuda T. and Adachi T. (1997) Characterization of a rice novel protein which binds to the α-globulin promoter. *Plant Mol. Biol.* **33**, 513-522.

Nijhawan A., Jain M., Tyagi A.K. and Khurana J.P. (2008) Genomic survey and gene expression analysis of the Basic Leucine Zipper transcription factor family in rice. *Plant Physiol.* **146**, 333–350.

Noma S., Kawaura K., Hayakawa K., Abe C., Tsuge N. and Ogihara Y. (2016) Comprehensive molecular characterization of the α/β-gliadin multigene family in hexaploid wheat. *Mol. Genet. Genomics* **291**, 65-77.

Norre F., Peyrot C., Garcia C., Rancé I., Drevet J., Theisen M. and Gruber, V. (2002) Powerful effect of an atypical bifactorial endosperm box from wheat HMWG-Dx5 promoter in maize endosperm. *Plant Mol. Biol.* **50**, 699-712.

Oñate L., Vicente-Carbajosa J., Lara P., Diaz I. and Carbonero P. (1999) Barley BLZ2, a seed-specific bZIP protein that interacts with BLZ1 in vivo and activates transcription from the GCN4-like motif of B-hordein promoters in barley endosperm. *J. Biol. Chem.* **274**, 9175-9182.

Onodera Y., Suzuki A., Wu C.Y., Washida H. and Takaiwa F. (2001) A rice functional transcriptional activator, RISBZ1, responsible for endosperm-specific expression of storage protein genes through GCN4 motif. *J. Biol. Chem.* **276**, 14139-14152.

Pfaffl M.W., Tichopad A., Prgomet C. and Neuvians T.P. (2004) Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper--Excel-based tool using pair-wise correlations. *Biotechnol. Lett.* **26**, 509-15.

Pysh L.D., Aukerman M.J. and Schmidt R.J. (1993) OHP1: a maize basic domain/leucine zipper protein that interacts with Opaque2. *Plant Cell* **5**, 227-236.

Plessis A., Ravel C., Bordes J., Balfourier F. and Martre, P. (2013) Association study of wheat grain protein composition reveals that gliadin and glutenin composition are trans-regulated by different chromosome regions. *J. Exp. Bot.* **64**, 3627–3644.

Ravel C., Fiquet S., Boudet J., Dardevet M., Vincent J., Merlino M., Michard R. and Martre P. (2014) Conserved cis-regulatory modules in promoters of genes encoding wheat high-molecular-weight glutenin subunits. *Front. Plant Sci.* **5**, 621.

Risacher T., Craze M., Bowden S., Paul W. and Barsby T. (2009) Highly efficient *agrobacterium*-mediated transformation of wheat via *in planta* inoculation. In *Transgenic Wheat, Barley and Oats, Methods in Molecular Biology* (Jones, H.D. and Shewry, P.R., eds), Humana Press, **478**, pp. 115–124.

Rubio-Somoza I., Martinez M., Abraham Z., Diaz I., and Carbonero P. (2006a) Ternary complex formation between HvMYBS3 and other factors involved in transcriptional control in barley seeds. *Plant J.* **47**, 269–281.

Rubio-Somoza I., Martinez M., Diaz I. and Carbonero P. (2006b) HvMCB1, a R1MYB transcription factor from barley with antagonistic regulatory functions during seed development and germination. *Plant J.* **45**, 17–30.

Schütze K., Harter K. and Chaban C. (2008) Post-translational regulation of plant bZIP factors. *Trends Plant Sci.* **13**, 247-55.

Shewry P.R., Tatham A.S. and Lazzeri P. (1997) Biotechnology of wheat quality. *J. Sci. Food Agric.* **73**, 397-406.

Shewry P.R., Tatham A.S. and Halford N.G. (2001) Nutritional control of storage protein synthesis in developing grain of wheat and barley. *Plant Growth Regul.* **34**, 105-111.

Shewry P.R. and Halford N.G. (2002) Cereal seed storage proteins: structures, properties and role in grain utilization. *J. Exp. Bot.* **53**, 947–958.

Sun F., Liu X., Wei Q., Liu J., Yang T., Jia L., Wang Y., Yan G. and He G. (2017) TaFUSCA3, a B3-Superfamily transcription factor gene in the wheat. *Front. Plant Sci.* **8**, 1133.

Thévenin J., Dubos C., Xu W., Le Gourrierec J., Kelemen Z., Charlot F., Nogué F., Lepiniec L. and Dubreucq B. (2012) A new system for fast and quantitative analysis of heterologous gene expression in plants. *New Phytol*. **193**, 504-12.

Thompson J.D., Desmond G., Higgins D.G. and Gibson T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673-4680.

Triboï E., Martre P. and Triboï-Blondel A.M. (2003) Environmentally-induced changes in protein composition in developing grains of wheat are related to changes in total protein content. *J. Exp. Bot.* **54**, 1731-1742.

Vandesompele J., De Preter K., Pattyn F., Poppe B., Van Roy N., De Paepe A. and Speleman F. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* **3**: research0034.

Van Herpen T.W., Riley M., Sparks C., Jones H.D., Gritsch C., Dekking E.H., Hamer R.J., Bosch D., Salentijn E.M., Smulders M.J., Shewry P.R. and Gilissen L.J. (2008) Detailed analysis of the expression of an alpha-gliadin promoter and the deposition of alpha-gliadin protein during wheat grain development. *Ann. Bot*. **102**, 331-342.

Verdier J. and Thompson R.D. (2008) Transcriptional regulation of storage protein synthesis during dicotyledon seed filling. *Plant Cell Physiol*. **49**, 1263-1271.

Vicente-Carbajosa J., Moose S.P., Parsons R.L. and Schmidt R.J. (1997) A maize zinc-finger protein binds the prolamin box in zein gene promoters and interacts with the basic leucine zipper transcriptional activator Opaque2. *Proc. Natl. Acad. Sci.* **94**, 7685-7690.

Vicente-Carbajosa J., Oñate L., Lara. P. Diaz I. and Carbonero P. (1998) Barley BLZ1: a bZIP transcriptional activator that interacts with endosperm-specific genes promoters. *Plant J.* **13**, 629-640.

Vincentz M., Bandeira-Kobarg C., Gauer L., Schlo P. and Leite A. (2003) Evolutionary pattern of angiosperm bZIP factors homologous to the maize Opaque2 regulatory protein. *J. Mol. Evol.* **56**, 105-116.

Vinson C., Myakishev M., Acharya A., Mir A.A., Moll J.R. and Bonovich M. (2002) Classification of human B-ZIP proteins based on dimerization properties. *Mol. Cell. Biol.* **22**, 6321-6335.

Weltmeier F., Ehlert A., Mayer C.S., Dietrich K., Wang X., Schütze K., Alonso R., Harter K., Vicente-Carbajosa J. and Dröge-Laser W. (2006) Combinatorial control of Arabidopsis proline dehydrogenase transcription by specific heterodimerisation of bZIP transcription factors. *EMBO J.* **25**, 3133-3143.

Wieser H., Antes S. and Seilmeier W. (1998) Quantitative determination of gluten protein types in wheat flour by reversed-phase high-performance liquid chromatography. *Cereal Chem.***75**, 644–650.

Xi D.M. and Zheng C.C. (2011). Transcriptional regulation of seed storage protein genes in *Arabidopsis* and cereals. *Seed Sci. Res.* **21**, 247-254.

Yang J., Ji C. and Wu Y. (2016) Divergent transactivation of maize storage protein zein genes by the transcription factors Opaque2 and OHPs. *Genetics* **204**, 581-591.

Zhang N., Qiao Z., Liang Z., Mei B., Xu Z. and Song R. (2012) *Zea mays* Taxilin protein negatively regulates Opaque-2 transcriptional activity by causing a change in its sub-cellular distribution. *PLoS One* **7**, e43822.

Zhang Z., Yang J. and Wu Y. (2015) Transcriptional regulation of zein gene expression in maize through the additive and synergistic action of opaque2, prolamine-box binding factor and O2 heterodimerizing proteins. *Plant Cell* **27**, 1162-1172.

Zörb C., Grover C., Steinfurth D. and Hermann Mühling K. (2010) Quantitative proteome analysis of wheat gluten as influenced by N and S nutrition. *Plant Soil* **327**, 225–234.

FIGURE LEGENDS

Figure 1. Multiple sequence alignment and phylogenetic tree analysis of the bZIP amino acid sequence of SHP, SPA (genome A), BLZ1 and BLZ2, and expression of *SHP* and *SPA* genes during grain development.

(a) Multiple amino acid sequence alignment. Asterisks indicate perfectly conserved amino acids, colons indicate strongly conserved amino acids, and periods indicate weakly conserved amino acids. The positions of hydrophobic leucine residues in the leucine zipper region are indicated in bold. (b) Phylogenetic tree drawn with the unweighted pair group method with arithmetic mean. (c) Quantitative RT-PCR measurements of expression of *SHP* and *SPA* in grains of the wheat cultivar NB1 from 300 to 700 °Cdays after anthesis. qRT-PCR was performed with generic primers to quantify the expression of the three homoeologous copies of *SHP* and *SPA*. Data are means ± 1 s.e. for $n = 4$ independent replicates.

Figure 2. Electrophoretic mobility shift assays (EMSA) of the recombinant SHP protein with oligonucleotides derived from the *GluB1-1* (a) and *GluD3* (b) promoters.

(a) EMSA of the recombinant SHP protein with 25-bp biotin-labelled GLM1 (-640), GLM2 (-619) and G-box (-269) probes derived from the *GluB1-1* promoter and their respective mutated versions *glm1*, *glm2* and *G-box*. (b) EMSA of the recombinant SHP protein with the 25-bp biotin-labelled GLM1 (-282) and GLM2 (-305) probes derived from the *GluD3* gene promoter and their respective mutated versions *glm1* and *glm2*. Numbers in brackets indicate the nucleotide position of the motif relative to the translation start site. Sequences of the oligonucleotides used as probes are shown with the cis-motifs in bold. Identical residues out of cis-motif are represented by dots and mutated nucleotides are shown in lower case.

Figure 3. Transient assays of *GluB1-1* promoter activity with SHP in wheat developing endosperms.

(a) Schematic representation of the reporter and effector constructs. The reporter constructs consisted of the *uidA* reporter gene (GUS) driven by 747 bp of the *GluB1-1* gene promoter (pGluB1-1) or mutated versions pGluB1-1* (both GLMs mutated), pGluB1-1** (G-box mutated) and pGluB1-1*** (G-box and both GLMs mutated). The nucleotide sequences are also shown. The effector constructs contained the complete cDNA of *SHP* under the control of the Ubiquitin promoter (pUbi), followed by the first intron of the *Ubi* gene (I-Ubi) and downstream the 3' nos terminator (nos).

(b) Transient expression assays in developing wheat endosperms co-bombarded with different molar ratios of the pGluB1-1 reporter and the pUbi-SHP effector.

(c) Transient expression assays in developing wheat endosperms co-bombarded with the pGluB1-1, pGluB1-1*, pGluB1-1** or pGluB1-1*** reporters with (1) or without (0) equimolar ratio of the pUbi-SHP effector. The insert shows significant differences (***, $P < 0.001$) between expression values for the pGluB1-1 and the mutated versions.

The number of bombarded endosperms from at least three independent particle bombardments varied between 43 and 119. In (b) and (c) asterisks above the data indicate significant differences between expression values for the reporter with the effector and the reporter without effector (***, $P < 0.001$; **, $P < 0.01$; *, $P < 0.05$). Median values are indicated by horizontal bars and mean values by crosses with corresponding values on the right.

Figure 4. Transient assays of *GluD3* gene promoter activity with SHP and SPA in wheat developing endosperms.

(a) Schematic representation of the reporter and effector constructs. The reporter constructs consisted of the *uidA* reporter gene (GUS) driven by 346 bp of the *GluD3* gene promoter (pGluD3). The nucleotide sequences are also shown. The effector constructs contained the complete cDNA of *SHP* or *SPA* under the control of the ubiquitin promoter (pUbi), followed by the first intron of the *Ubi* gene (I-Ubi) and downstream the 3' nos terminator (nos).

(b) Transient expression assays in developing wheat endosperms co-bombarded with the pGluD3 reporter and equimolar ratios of pUbi-SHP or pUbi-SPA effectors.

The number of bombarded endosperms from at least 3 independent particle bombardments varied between 17 and 82. In (b), asterisks indicate significant differences between expression values for the reporter with the effector and the reporter without the effector (***, $P < 0.001$). Medians are indicated by horizontal bars and means by a cross with the corresponding value to the right.

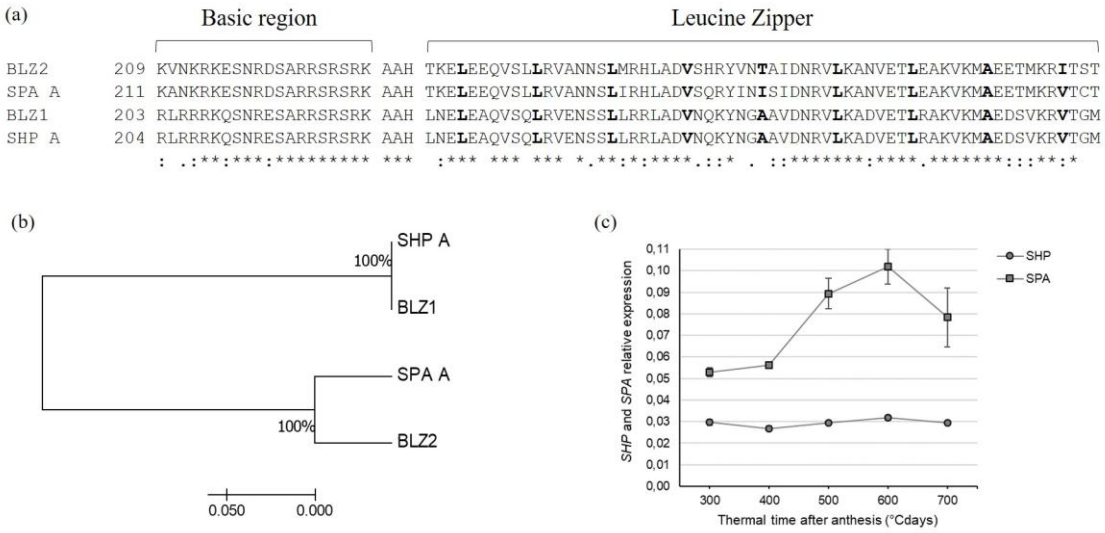
Figure 5. *SHP* and *SPA* expression in grains at 500 °Cdays after anthesis from *SHP* overexpressing wheat (OE) derived from two independent transgene insertion events and from the corresponding null segregants (NS) grown in the greenhouse with low (N-) and high (N+) nitrogen supply. qRT-PCR was performed with generic primers to quantify the expression of the three homoeologous copies of *SHP* and *SPA*. Results for the two

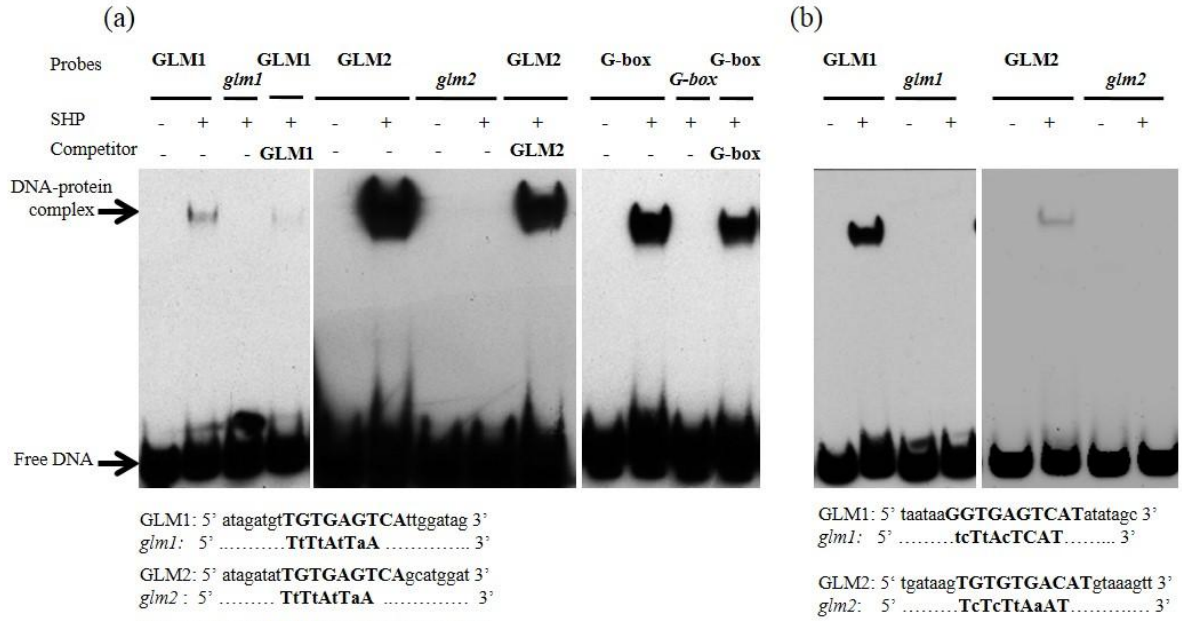
independent OE progenies were similar and were pooled for analysis. Data are means ± 1 s.e. for n = 8 and n = 4 independent replicates for the OE and NS lines, respectively. Asterisks and ns indicate significant (***, $P < 0.001$) and no significant differences from NS values.

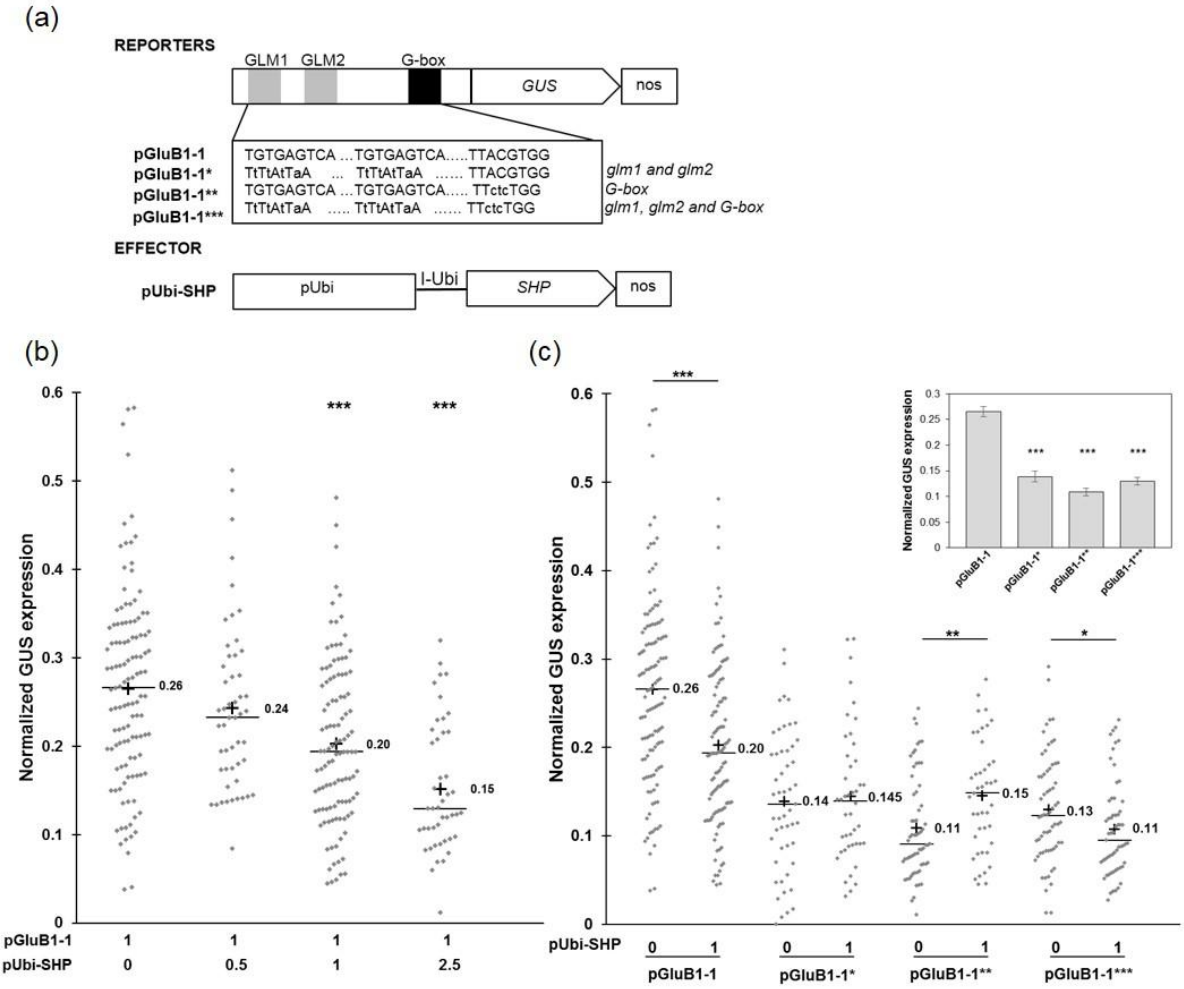
Figure 6. Grain storage protein content of mature grains from *SHP* overexpressing wheat (OE) derived from two independent transgene insertion events and from their corresponding null segregants (NS) grown in the greenhouse with low (N-) and high (N+) nitrogen supply. Results for the two independent OE progeny were similar and were pooled for analysis. Data are means ± 1 s.e. for n = 8 and n = 4 independent replicates for the OE and NS lines, respectively. Asterisks indicate significant differences from NS values (***, $P < 0.001$; **, $P < 0.01$; *, $P < 0.05$).

Figure 7. Illustration of the passive repression mechanism of SHP on the *GluB1-1* gene promoter.

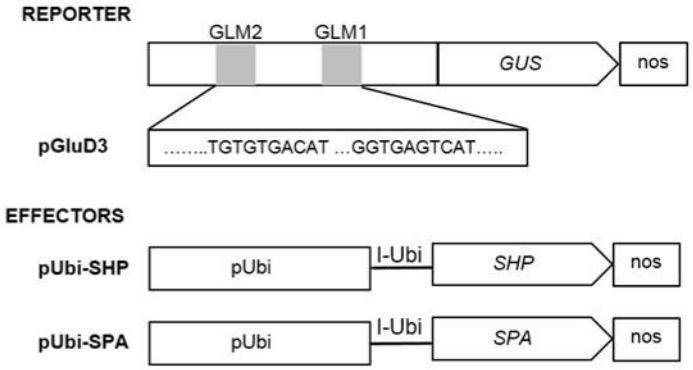
As SPA is more abundant than SHP, all bZIP motifs can be occupied by SPA homodimers. On the contrary, SHP homodimers bind all cis-motifs in grain overexpressing *SHP* gene. SHP prevents SPA binding and activating activity.







(a)



(b)